

Članci i rasprave

UDK: 371.275:371.279.7:811.163.42'243

Izlaganje sa znanstvenoga skupa

Prihvaćeno za tisak: 18. prosinca 2008.

Kako izraditi jezični test

Ina Ferbežar i Vesna Požgaj-Hadžić

Ljubljana, Slovenija

Mjerenje jezičnoga znanja u sklopu učenja/poučavanja stranih/drugih jezika jedna je od glavnih tema evropske jezične politike. Zasad su najsigurniji (formalni) mjerni instrumenti jezični testovi koji moraju biti izrađeni po određenim kriterijima. Ako se govori o komunikacijskoj metodi učenja i mjerenja znanja, mora se naglasiti da je temeljni koncept izrade testova i provođenja testiranja autentičnost — i to ne samo autentičnost tekstova već i autentičnost testnih situacija. U radu se predstavljaju različite karakteristike testova (od valjanosti kao najznačajnije do objektivnosti, osjetljivosti itd.) i različite namjene testiranja (od neformalnoga provjeravanja uspješnosti za vrijeme učenja jezika do formalnoga javnog sustava certificiranja jezičnoga znanja). Posebna se pozornost posvećuje mnogim otvorenim problemima pri izradi testova, kao npr. izboru tekstova, različitim tipovima zadataka (i provjeravati li njima eksplicitno gramatiku), ocjenjivanju, odnosno mjerilima za ocjenjivanje i vrednovanje rezultata itd. Na primjerima testova za hrvatski jezik ukazuje se na brojne "klopke" pri izradi testnih zadataka o kojima moraju voditi računa sastavljači testova. Zaključuje se da izrada testova mora biti dobro promišljena jer rezultati testa ponekad mogu i sudbinski odlučivati o pojedincima. Zato sastavljači testova i njegovi izvođači moraju biti svjesni svoje odgovornosti.

0. Uvod

Prije početka ocijenite govornika u tekstu u (1).

- (1) A Dobar dan.
B Dobar dan.
A Pa ja ću...

- B Kako Vam mogu pomoći?
- A Pa da, provali su mi u auto d..., danas sam, danas-sutra kad sam htio da odem na posao našao sam da ... s... da su mi provali u auto.
- B M-hm... kada je to točno bilo?
- A Ma ne znam. Ja sam juče ... uvečer sam parkirao auto ispred mog stana, a danas-sutra kad sam htio da idem bio je ... bio je osim ... bio je ulomljen u auto.
- B M-hm ... emm ... gdje ste ga točno parkirali?
- A Ma tamo gdje ga vjedno ... tamo gdje ga uvijek parkiram.
- B Emm ... možda ste ga parkirali na nekom drugom mjestu?
- A Ne, ne, ne, ne. Pa tamo gdje ga uvijek ...
- B Sigurno ga uvijek na tom mjestu parkirate.
- A Da, da, uvijek, uvijek. (...)

Kod zadatka koji vam postavljamo sigurno biste (nas) željeli pitati kao u (2).

- (2) — Koga da ocjenjujem? (Najvjerojatnije ste ocjenjivali stranoga govornika — profesionalna deformacija.)
- Što da uopće ocjenjujem? (Najvjerojatnije ste — iz navedenoga razloga — ocjenjivali njegovu jezičnu kompetenciju.)
- Po kojim mjerilima? Kakav da dakle “metar” upotrijebim? (Vjerojatno ste se ograničili na ocjenjivačke postupke koje i inače upotrebljavate.)
- Što je svrha mogega ocjenjivanja? Za što će se upotrijebiti rezultati? Hoće li se govornici uspoređivati? Kako? Što će značiti razlika između pojedinih ocjena?

Odgovori na ta pitanja sigurno su temelj svakoga mjerenja, bilo jezičnoga, bilo nekoga drugoga. Zato ih moramo dobro poznavati i prije nego što se prihvatimo izrade bilo kakvoga testa. U polazište razmišljanja o jezičnom testiranju staviti ćemo neposrednoga korisnika testa — u našem slučaju stranoga govornika čiju (više ili manje) ograničenu jezičnu kompetenciju želimo izmjeriti.

1. Za početak: malo teorije

Kada započinjemo mjerenje, moramo najprije postaviti teoretski model onoga što namjeravamo mjeriti. U našem slučaju to će biti model jezične kompetencije, odnosno ako upotrijebimo općenitiji izraz, jezičnoga znanja (usp. Bachman i Palmer 1996). Međutim, pravilno definirati jezično znanje vrlo je složena zadaća: riječ je o jako zamršenoj i kompleksnoj pojavi. Najutjecajniji modeli — barem oni kojima se jezična kompetencija pokušala definirati upravo za potrebe mjerenja (npr. Bachman 1990, Bachman i Palmer

1996) — predstavljaju je kao sastavljenu kompetenciju. Tako je u Bachman-Palmerovu modelu jezično znanje (engl. *language knowledge*) sastavljeno od organizacijskoga znanja, koje nadalje čine gramatičko (leksik, sintaksa te pravogovor i pravopis) i tekstno znanje (kohezija i retorička ili konverzijska organizacija), te od pragmatičnoga znanja, koje je sastavljeno od funkcionalnoga i sociolingvističkoga znanja. Sve te dijelove jezičnoga znanja usmjerava strateška kompetencija (Bachman i Palmer 1996: 67–75). Autori su jezično znanje tako “razdrobljeno” predočavali zato što ga je kao takvo najlakše mjeriti — lakše je mjeriti njegove pojedine segmente, pa i nepotpune. Kad bismo ga željeli izmjeriti u čitavoj njegovoj cjelovitosti, morali bismo pratiti korisnika jezika na njegovim jezičnim putovima i dulje vrijeme pratiti njegovu jezično ponašanje. Međutim, to bi bilo vrlo neekonomično, iako bi bilo najautentičnije — ali o autentičnosti ćemo nešto kasnije. Tako nam ostaje “parcijalno” mjerenje pojedinih elemenata jezične kompetencije ograničenim instrumentima — jezičnim testovima.

Jezična kompetencija obično se definira kao apstraktni sustav pravila i različitih kompetencija (npr. gramatičke, leksičke, sociokulturne itd.) i bez konkretnih okolnosti samo je potencijalna kompetencija govornika. To znači da njezinu realizaciju možemo pratiti tek u konkretnoj situaciji koja se očituje jezičnom uporabom, odnosno jezičnom produkcijom koju stručno nazivamo performancija. Ona je, naime, “vidljiva” te zato i mjerljiva; tek se iz “konkretnu” performancije može zaključiti kakva je “apstraktna” sposobnost, odnosno kompetencija (Ferbežar 1999: 423–424).

Naravno da se otvara pitanje koja jezična kompetencija treba biti polazište za naš model. Da li da polazište bude ideal(izira)ni domaći govornik kojega je generativno jezikoslovlje (a poslije njega i strukturalno) smatralo referentnim mjerilom za jezičnu produkciju (Spillner 1995: 13)? Ili da, kao što predlažu novije teorije, polazište za ocjenu stupnjeva jezične kompetencije bude spretni strani korisnik jezika?

Bez obzira na to na koju se od teorija oslanjamo, možemo se složiti da je jezična kompetencija promjenjiva: učenjem i uporabom može se odgovarajuće raširiti, a s druge strane sklona je i nezadrživim procesima smanjivanja kada jezik prestanemo učiti i ne upotrebljavamo ga. Ukratko, dinamična je i stupnjevita. Pitanja stupnjevitosti jezične kompetencije u prvome jeziku rijetko se postavljaju — u najboljem slučaju kod istraživanja funkcionalne pismenosti; sama po sebi razumljivom čini se kod jezične kompetencije stranih korisnika nekoga jezika. To ističemo zato što se dinamična priroda jezičnoga znanja u postupcima jezičnoga testiranja nekako gubi. Činjenica je, naime, da jezični test “zamrzava” jezično znanje na onome stupnju za koji je test namijenjen (ili ispod njega) — ovisno o tome kakav je rezultat pojedinac postigao na testu. Na to sigurno snažno utječu i drugi čimbenici koji nisu (nužno) povezani s njegovim jezičnim znanjem (npr. trema na tes-

tu, vrijeme dana, odnos prema stvarima itd.). U testologiji govorimo o tzv. testnoj performanciji (Ferbežar 1999) koja osim spomenutoga odražava i pojedinačnu spretnost rješavanja testova, slično rješavanju križaljki, slagalica, sudokua.

Rekli smo već da je za sada najpraktičniji instrument za mjerenje i ocjenu jezične kompetencije jezični test. Ovdje nema ni vremena ni mjesta za detaljno upuštanje u to što je (jezični) test. Pojednostavljeno možemo reći da je riječ o standardiziranom postupku koji izaziva određenu aktivnost, a ona odražava stupanj mjerene kompetencije (u našem slučaju jezične) koji bi test trebao barem približno objektivno izmjeriti, i to tako da se pokažu razlike između osoba koje testiramo. Test je dakle nekakvo pitanje — ako je pravilno postavljeno, dat će i odgovarajući, pouzdan odgovor. U testologiji govorimo o valjanosti (što mjerimo), objektivnosti (kako mjerimo), osjetljivosti (razlike između osoba) i pouzdanosti testa. Pouzdanost je povezana s testom i njegovim bodovanjem: test je pouzdan kada ista osoba daje jednak odgovor na isto pitanje (uz pretpostavku da se u tom vremenu njezina kompetencija koju testiramo ne mijenja) i da se njezin odgovor uvijek ocjenjuje jednako — bez obzira na to tko (i što) boduje, odnosno ocjenjuje test (Ferbežar i Bajec 2004).

Problematika jezičnoga testiranja, a posebno ocjenjivanja, vrlo je složena; o njoj je napisano mnoštvo knjiga i zato je na ovako ograničenom prostoru nećemo moći detaljnije predstaviti. Zato ćemo radije kod konkretnih primjera otvoriti nekoliko pitanja i ponuditi ih na razmišljanje te upozoriti na neke zamke koje “postavlja” jezično testiranje. Međutim, prije nego što krenemo na konkretne primjere, još nekoliko riječi o metodici poučavanja i testiranja.

Od 70. godina prošloga stoljeća u jezičnom poučavanju i, posljedično, u testiranju uvriježila se tzv. komunikacijska metoda koja je učenje od znanja o jeziku pomaknula prema učenju za odgovarajuće jezično sporazumijevanje u različitim situacijama. U jezični razred i jezične testove donijela je simulaciju realne jezične uporabe. Iz komunikacijske metode posljednjih se godina razvila tzv. *task-based* metoda, koja usmjerava prema odgovarajućem jezičnom ponašanju zadacima koji izazivaju “pravu” (ne simuliranu) interakciju među onima koji uče jezik. Ta metoda koja usmjerava prema jezičnome ponašanju potvrdila se u jezičnim razredima i jezičnom testiranju. Osobito se upotrebljava pri poučavanju i testiranju djece u predškolsko doba, a i kod odraslih s posebnim potrebama (npr. nepismenih). S njome je tijesno povezan koncept autentičnosti. To je vrlo važan koncept koji se odnosi na tekstove uključene u nastavu jezika te na situacije, odnosno jezične zadatke koji prate te tekstove i reakcije na te zadatke. Kada govorimo o autentičnosti tekstova, ne mislimo samo na tzv. originalne tekstove. Zapravo je većina tekstova koji se uključuju u jezične testove na jedan ili drugi način priređena

ili čak napisana posebno za testove. Pitanje originalnosti tu treba raširiti na pitanje relevantnosti tekstova za ciljnu publiku. Ako želimo izazvati autentičnu jezičnu uporabu kakva je karakteristična i za netestne komunikacijske situacije, moramo s jedne strane uzeti u obzir tekstove relevantne za ciljnu publiku. S druge strane, zadaci kojima mjerimo tu jezičnu uporabu moraju biti pripremljeni tako da izazovu autentičnu reakciju. To u praksi znači da testirane osobe zadatkom treba staviti u takvu situaciju u kojoj će imati mogućnost odazvati se odgovorom koji je za njih tipičan, uobičajen, dakle autentičan, svejedno je li riječ o odgovorima riječima ili drukčijima (Ferbežar i Bajec 2004).

2. Planiranje testa — specifikacija

Prije nego što započnemo izradu testa, moramo pripremiti odgovarajuću specifikaciju testa, koju smo donekle najavili već u uvodu. U specifikaciji trebaju biti detaljno definirani sljedeći dopunski i prepleteni parametri (Ferbežar i Bajec 2004: 166–168).

Svrha testa. Svrhe su jezičnih testova različite, u grubo ih možemo podijeliti na one testove koji su namijenjeni za jezikoslovna istraživanja, odlučivanje o jezičnim programima i odlučivanje o ljudima. U jezičnom testiranju najviše se upotrebljavaju osobito potonji i smatraju se više-manje sudbinskima. Oni odlučuju o tome hoće li pojedinac biti prihvaćen ili odbijen (prijemni testovi), kako će biti uvršten (testovi uvrštavanja), hoće li napredovati ili ne (testovi napredovanja), hoće li uspješno završiti neki nastavni proces ili ne (testovi uspješnosti). Nabrojeni su testovi više-manje povezani s nastavnim procesom, a postoje i testovi koji provjeravaju jezičnu kompetenciju bez takve povezanosti: riječ je o vanjskim provjeravanjima koja proizlaze iz službenih potreba (npr. za obavljanje posla, upis na sveučilište, dobivanje državljanstva itd.) ili nekih drugih specifičnih potreba (tzv. testovi za posebne svrhe) i zato može imati vrlo široke društvene posljedice. One su posebno značajne.

Ciljna publika. Sa svrhom testa tijesno je povezana i ciljna publika. Naravno da nije svejedno hoće li naše testirane osobe biti djeca ili odrasli, jesu li početnici ili usavršavaju svoje znanje, jesu li izvorni ili strani govornici nekoga jezika, je li riječ o skupini specijalista kojima je potreban poseban test. Svaka od nabrojenih skupina zahtijeva drukčiji test.

Svrha testa i ciljna skupina naravno definiraju i naglasak testiranja, to je ono što *mjeri naš test*. U specifikaciji moramo predvidjeti koje ćemo kompetencije mjeriti njime — opću jezičnu kompetenciju, samo neki od njezinih

elemenata ili određene specifične kompetencije za posebne potrebe. Dakle, moramo točno definirati konstrukt. Konstrukt je u testologiji hipotetična kompetencija koja se ne može neposredno pratiti. Takav je konstrukt u našem slučaju jezična kompetencija sa svim svojim elementima (Ferbežar i sur. 2004: 130).

Dijelovi testa i njihova vrijednost. Taj je parametar tijesno povezan s naglaskom testiranja i od ispitivača zahtijeva definiciju toga koliko će dijelova i koji će dijelovi sastavljati test i koliko će ti dijelovi, ako ih bude više, biti “teški”. Hoće li donijeti jednake udjele zajedničkoj ocjeni ili će neki od dijelova biti “teži” te time važniji? S dijelovima testa povezano je naravno i vremensko ograničenje koje moramo barem približno predvidjeti. Zato je vrlo važno da testovi budu pretestirani na sličnoj populaciji kakva je naša ciljna. Jezične testove koji mjere opću jezičnu kompetenciju obično čine dijelovi koji se odnose na četiri komunikacijske aktivnosti: receptivne aktivnosti slušanje i čitanje te produktivne govorenje i pisanje (s obzirom na specifičnost pojedinoga testa gramatika može biti njegov poseban dio, a može biti uključena u produktivne dijelove testa).

Tekstovi i druga građa. Za test nije tako važno jesu li tekstovi na kojima se zasnivaju testni zadaci originalni, priređeni ili posebno napisani. Mnogo je važnija njihova relevantnost za ciljnu populaciju. Osobe koje ćemo testirati morale bi se barem približno identificirati s tekstom koji moraju pročitati jer će ga tako mnogo lakše i bolje shvatiti. Možda će tekst pratiti i neki neverbalni poticaji, kao što su slike, grafikoni, tablice, kod slušanja šumovi, glazba i dr. Slikovni materijali osobito su važni kod testiranja djece. Kod tekstova za slušanje koje uključujemo u testove slušanja moramo paziti da oni budu namijenjeni slušanju. Ako je tekst u polazištu namijenjen za čitanje, treba ga za slušanje dobro prirediti. Poseban je zahtjev kod izbora materijala nepristranost: tekstovi i dodatni materijali nikad ne daju prednosti nekoj od skupina ispitanika niti ih diskriminiraju ili čak vrijeđaju. Zato vrlo oprezno sa šalama i političkim temama.

Tipovi zadataka tijesno su povezani s tekstovima (ako su naravno zadaci vezani uz tekst — usp. u nastavku). Često već sam tekst nudi i tip zadatka. Za test je važno da za mjerenje pojedinih potkompetencija (npr. čitanja, slušanja itd.) upotrijebimo različite tipove zadataka koje naši ispitanici poznaju. Nekim ljudima bolje leže izborni tipovi zadataka, a drugi bolje rješavaju otvorene tipove. U testovima bi to trebalo imati u vidu i ponuditi “svakome nešto”.

*Mjerila bodovanja i ocjenjivanja.*¹ Ocjenjivačka mjerila moraju biti točno određena unaprijed i mora ih poznavati i naša testna populacija. Kod zadataka objektivnoga tipa to znači da moramo sve odgovore predvidjeti unaprijed, i to ne samo pravilne već i moguće i vjerojatne, a oni se moraju i jasno bodovati. Tako npr. kod receptivnih zadataka zapravo ocjenjujemo razumijevanje i ne odbijamo bodove za gramatičke pogreške u odgovoru (ako je npr. riječ o otvorenim pitanjima). Kod govorne i pisane produkcije mora biti jasno ocjenjivaču i ispitaniku što će se ocjenjivati i kako (npr. jezična pravilnost, sadržaj, leksik. . .). Ako testirane osobe znaju što ih na testu čeka, tome će posvećivati veću pozornost. To znači izradu jasnih ocjenjivačkih ljestvica, analitičkih i/ili cjelovitih, već u specifikaciji. Ocjenjivanje po takvim ljestvicama zahtijeva odgovarajuću stručnu osposobljenost ocjenjivača koji se na ljestvice moraju odgovarajuće “kalibrirati” da ih mogu tumačiti barem približno jednako. Samo tako možemo postići pouzdanost ocjenjivanja. Idealno je ako produktivne dijelove testa mogu ocjenjivati dva ocjenjivača — jedan po analitičkoj ljestvici (po pojedinim parametrima, kao što su sadržaj, jezična pravilnost, interakcija, stil, leksik itd.), a drugi holistički (po dojmu).

3. Zamke u izradi testnih zadataka (primjeri . . .)

U ovome ćemo poglavlju pogledati neke tipove zadataka, predstaviti njihove prednosti i mane te upozoriti na što autor testa mora paziti pri pripremi testnih zadataka.

3.1. Zadatci bez teksta

Testne zadatke ugrubo možemo podijeliti na zadatke bez teksta i zadatke koji proizlaze iz teksta. Prvi su namijenjeni uglavnom provjeravanju gramatičkih struktura i nisu komunikacijski, a još manje autentični. Ograničavaju se na znanje o jeziku, a ne znanje jezika i kao takvi ne mjere jezičnu uporabu u konkretnoj komunikacijskoj situaciji, primjeri su u (3) i (4).

(3) *Oblikujte rečenice u prošlom (perfektu) i budućem vremenu (futura).*

Ovdje sjedim od dva sata.

Želite otputovati iz grada.

¹U ovome dijelu razlikujemo bodovanje, koje se odnosi na objektivne dijelove testa i provodi se na temelju tzv. bodovnih šablona, i ocjenjivanje, koje se odnosi na produktivne dijelove testa i vezano je za subjektivnu ocjenu odgovarajuće izvježbanih ocjenjivača.

(4) *Zaokružite slovo ispred točno napisane rečenice.*

- a. Nekoliko djece igraju se u parku.
- b. Nekoliko djeca igra se u parku.
- c. Nekoliko djece igra se u parku.

Tzv. zadaci nepovezanih jedinica (engl. *discrete item*) sigurno su upotrebljivi za mjerenje određenoga segmenta jezične kompetencije. Međutim, budući da nisu autentični, u jezičnim se testovima napuštaju na račun kontekstualiziranih zadataka, odnosno zadataka koji proizlaze iz teksta.

3.2. Zadaci izbora

Ima više zadataka izbornoga tipa. Jedan je među njima izbor između dviju mogućnosti — pravilne i nepravilne, kao što se vidi u primjeru (5).

(5) *Pažljivo slušajte razgovor,² a zatim za navedene tvrdnje odgovorite jesu li točne ili nisu. (2 boda za svaki točan odgovor. Ukupno 8 bodova.) Podvucite točan odgovor.*

- 1. Prilikom čekiranja, odnosno prijavljivanja Marko treba predati prtljagu. *točno netočno*
- 2. Kad se objavi početak ukrcavanja, Marko treba proći kroz kontrolu putovnica. *točno netočno*
- 3. Prilikom ukrcavanja Marko ne treba pokazati stjuardesi voznu kartu. *točno netočno*
- 4. Ako se izgubi, neka pita osoblje koje radi u zračnoj luci da mu pomogne jer su veoma neljubazni i neugodni. *točno netočno*

Iako zadatak omogućuje veliku mjeru pogađanja, smatra se jednim od najautentičnijih i često se upotrebljava pogotovo za testiranje razumijevanja govorenih tekstova (slušanja), i to razumijevanja glavnih dijelova teksta. Međutim, zadatak u (5) uzrokuje poteškoće. Prva je od njih da bi se zadatak mogao riješiti bez čitanja teksta jer navodi neke opće tvrdnje.

Ako imate iskustvo sa zračnom lukom, riješit ćete ga bez problema. Navođenje općih tvrdnji i inače je problem toga tipa zadataka pa je zato jedno od mogućih rješenja problema da ga nakon pripreme damo nekome riješiti bez teksta. Još dva konkretna problema: treća tvrdnja u negativnom je obliku. Negacija može zavesti ispitanika jer se, barem na hrvatskom jeziku,

²Da biste riješili ovaj zadatak, tekst za slušanje nije potreban.

s tvrdnjom može složiti na dva načina, *da, ne treba*, ili *ne, stvarno ne treba*. Zato je dobro da ispitanici izabiru između odgovora *točno* i *netočno*, a ne između odgovora *da* i *ne*.

Problematična je i četvrta tvrdnja — kao prvo zbog svoje dužine. Riječ je o zadatku za slušanje kojim ne provjeravamo i čitanje (iako se to naravno ne može posve izbjeći). I kao drugo zbog potpune nelogičnosti tvrdnje. Možemo zaključiti da u takvom zadatku tvrdnje moraju biti vrlo precizno postavljene; ako su tvrdnje u negativnom obliku, predlažemo da NE BUDU dvosmislene i da se istaknu velikim slovima, kosim tiskom i sl.

Da bi se smanjio stupanj pogađanja, testolozi su kod ovoga tipa zadatka uveli treću mogućnost — *nije u tekstu*. Međutim ta je dodatna mogućnost problematična jer se iz svih tekstova može štošta zaključiti — pa i onda kad nije eksplicitno izraženo, ali možda usprkos tomu može biti samo po sebi razumljivo. I u slovenskim testovima³ zadatak s trećom mogućnošću pokazao se statistički neprimjerenim.

Zadatak u (5) donosi još jednu poteškoću: bodovanje. Za svaki odgovor predviđena su dva boda, što je neobično. Pogađanje se s jedne strane nagrađuje, a s druge strane naši se ispitanici za pogrešan odgovor jako kašnjavaju. Obično u testovima objektivnoga, odnosno zatvorenoga tipa, kakav je prikazan u (5), pravilna rješenja bodujemo jednim bodom. Takav način bodovanja vrijedi i kod otvorenih pitanja, koja se inače smatraju poluotvorenim tipom zadataka. I kod otvorenih pitanja treba vrijediti pravilo da jedna informacija donosi jedan bod.

Drugi izborni tip, nekako klasičan, jest izbor između triju ili više mogućnosti, kao u (6).

- (6) *Izaberite pravilan oblik, kao što je prikazano u primjeru. Od četiri oblika pravilan je samo jedan.*

Primjer:

Je li Beyonce Knowles (26) željela/želio/želi/želeo podržati ili uvrijediti Kate Moss odjenuvši majicu na koju/kojoj/kojom/koje ispod njegove/njezine/njezinu/svoje fotografije piše...

Primjer u (6) najuobičajeniji je primjer izbornoga tipa zadatka koji se upotrebljava za provjeravanje poznavanja gramatičkih struktura, te razumijevanja čitanja u primjeru u (7), pri čemu se može provjeravati razumijevanje glavnih dijelova teksta i detaljnosti. Ako se primjeru u (5) prigovaralo da se kod njega može pogađati, pogađanje je s četiri mogućnosti koje su

³Za ilustraciju dalje u tekstu navodimo iskustva u testiranju slovenskog jezika kao drugog/stranog.

na raspolaganju za izbor smanjeno na minimum. Treba međutim reći da je za sastavljače vrlo teško naći tri više-manje jednakovrijedna distraktora. Distraktori su nepravilni odgovori koji moraju biti dovoljno vjerojatni da svakoga od njih izabere barem pokoji ispitanik (Ferbežar i sur. 2004). Ako ih ne izabere nitko, zapravo je autorov (i ispitanikov) posao bilo uzaludan. Utvrđeno je da je taj tip zadataka posve pouzdan i onda kada treba izabirati između samo tri mogućnosti. U zadatku u (6) problemi su drugdje: ne samo što u jednoj rečenici ima puno nepoznanica već ispitanik mora znati jesu li Beyonce i Kate muškarci ili žene budući da se samo tako može točno riješiti prvo i treće pitanje. To više nije jezično znanje. Osim toga, prva stavka navodi mogućnost koja u hrvatskom jeziku ne postoji (*želeo*). Takve distraktore treba izbjegavati: zadatak ionako zavodi (s tri nepravilne mogućnosti) te zato nije u redu da zavodimo još i oblicima kojih nema u jeziku. Slovenski su testovi na primjer pokazali da takvi distraktori ne razlikuju između dobrih i loših ispitanika. Upravo suprotno: njih izabiru ispitanici koji su na testu inače ocijenjeni bolje.

Usprkos brojnim negativnim svojstvima zadatak se često uključuje u jezične testove — ispitanici naime točno znaju što moraju učiniti, čak i da u zadatku u (6) to nije posebno objašnjeno, bez obzira na to kako će to učiniti (postoje različite strategije: najuobičajenija je izdvajanje najnevjerojatnijih mogućnosti). Nastavimo sa sličnim primjerom, u (7).

(7) *Pročitajte tekst i izaberite pravilan odgovor kao što je prikazano u primjeru.*

- Konobar!
- Izvolite!
- Dajte mi jednu domaću rakiju i donesite mi jelovnik. Želim nešto pojesti.
- Izvolite. Za predjelo vam mogu preporučiti dobru pileću juhu ili palentu, a za glavno jelo odrezak s pohanom sirom i prženim krumpirićima kao prilogom.

Primjer: U ovom dijalogu jela preporučuje...

| | | | |
|----------|-------|----------------|-------|
| liječnik | kuhar | <i>konobar</i> | pekar |
|----------|-------|----------------|-------|

1. Dijalog se odvija...

| | | | |
|-----------|------------|-----------|-------------|
| u bolnici | na tržnici | u mesnici | u restoranu |
|-----------|------------|-----------|-------------|

2. Gost za aperitiv želi...

| | | | |
|------------------|--------|-------------|---------------|
| H ₂ O | rakiju | domaće vino | svijetlo pivo |
|------------------|--------|-------------|---------------|

Prva primjedba u zadatku u (7) odnosi se na tip teksta. Izborni tip zadatka uz tekst posve odgovara, međutim gornji tekst, budući da je dijaloški,

prikladniji je za slušanje nego za čitanje (ovdje neuljudnost gosta nećemo ni komentirati). Još se jedanput pozivamo na autentičnost situacije i zadatka: iako je možda riječ o originalnom tekstu, on ipak nije autentičan jer su dijalozi u restoranu uvijek govoreni — osim možda u dramskim djelima. Osim toga, riješeni primjer odgovara na sljedeće pitanje: ako jela preporučuje konobar, dijalog se najvjerojatnije događa u restoranu. Na to pitanje inače odgovara i drugo pitanje o aperitivu koji gost naručuje... U testologiji postoji pravilo da pitanja ne proizlaze jedno iz drugoga, dakle pravilno rješenje jednoga pitanja ne smije biti uvjet za pravilno rješavanje nekoga od sljedećih pitanja.

Kod drugoga pitanja smeta također simbol za vodu (H_2O): budući da je jedina simbolom zapisana mogućnost, među ostalima ima posebnu vrijednost — možda će upravo zato biti privlačnija ili je pak — što je vjerojatnije — neće nitko izabrati. Osim toga poznavanje simbola spada u drugu vrstu znanja, ne jezično. Ako želimo provjeriti takvo znanje, to je naravno nešto drugo. Međutim, onda bismo u odgovarajući simbol trebali pretvoriti i preostale tri mogućnosti.

3.3. Zadaci povezivanja

Pogledajmo sljedeći primjer u (8).

(8) *Dolje navedene osobe željele bi riješiti svoj stambeni problem. Pažljivo pročitajte oglase te povežite oglase i situacije: pored brojke upišite slovo koje se nalazi uz onaj oglas za koji mislite da odgovara navedenoj situaciji. Pazite, četiri su oglasa suvišna! (1 bod)*

- a) Izdajem garsonijeru u predgrađu Zagreba. Cijena je 130 eura plus režije.
- b) Izdajem sobu s upotrebom kupaoice i kuhinje u Novom Zagrebu. Cijena je povoljna.
- c) Izdajem dvokrevetnu sobu na Črnomercu uz mogućnost upotrebe telefona.
- d) Izdajem novi dvosobni nenamješten stan u Zagrebu. Cijena je izuzetno privlačna.
- e) Povoljno izdajem sobu s upotrebom kupaoice i kuhinje u centru Zagreba.
- f) Izdajemo dvosobni stan u Prečkom s centralnim grijanjem i telefonom. Stan se nalazi na 8. katu.

- 1. Marko je upisao fakultet u Zagrebu i mora pronaći sobu, ali s time da može koristiti kupaoicu i kuhinju u tom stanu.
- 2. Marija i Ivan mladi su braćni par bez djece koji traži dvosobni stan u Zagrebu. Nije im važno gdje se stan nalazi, ali je obavezno da ima centralno grijanje.

Primjer u (8) pripada u zadatke povezivanja kod kojih ispitanici povezuju dijelove teksta s obzirom na ključne informacije u njima. Zadatak se upotrebljava za provjeravanje razumijevanja zapisanih tekstova i smatra se prilično autentičnim. Međutim, zadatak je pravilno pripremljen samo ako su moguća povezivanja posve nedvosmislena i jasna, što za primjer u (8) ne bismo mogli tvrditi. Moguće je nekoliko povezivanja, a osim toga zadatak zahtijeva određenu mjeru geografskoga znanja. Dobro je da ima pokoja suvišna mogućnost, ali neka ih ne bude previše.

Zadatak u (8) problematičan je i zbog svoje upute. Zbog preciznosti i time suvišnih riječi zadatak ima upravo suprotan učinak: djeluje vrlo komplicirano, a on je zapravo jednostavan: treba povezati slova i brojeke. Dobra praksa jezičnoga testiranja zahtijeva da uputa bude jasna i jednoznačna, napisana tako da ne bude teže razumljiva od zadatka.

Ima više vrsta zadataka povezivanja — njima se može provjeravati poznavanje leksika, gramatičkih struktura i sl.

3.4. Zadaci dopunjavanja

(9) *Poslušajte tekst⁴ i popunite tablicu informacijama:*

| | |
|------------------------|-------|
| Država: | _____ |
| Nadmorska visina: | _____ |
| Značenje: | _____ |
| Udaljenost od Zagreba: | _____ |

Popunjavanje tablica glavnim informacijama u zapisanom ili govorenom tekstu jedan je od primjera dobre prakse. Prije svega zato što se ograničava na jednu od aktivnosti (ili čitanje ili slušanje) i istodobno ne zahtijeva previše pisanja. Rekli smo doduše da se takvo prekrivanje ne može posve izbjeći, a gornji primjer ograničava je na minimum.

Nešto je drukčiji tip zadatka dopunjavanja sljedeći, kao u (10).

(10) *Popunite tekst.*

Upravo sam završavala posao kad je iznenada počela _____ kiša. Netko je _____ moj kišobran. Ne znam _____. Moje kolegice su _____

⁴Tekst za slušanje: "Desetak kilometara od centra Zagreba, u sklopu parka prirode Medvednica, nalazi se njezin najviši vrh, Sljeme. Od pamtivijeka zagrebačko izletište, skijalište, sanjkaliste te isto tako pogodno za planinarenje, svjež zrak i uživanje u prirodi. Najviši vrh Zagrebačke gore smješten je na visini od 1033 m. Sljeme je od 2005. g. ponosni dom FIS svjetskog kupa za skijašice, a uskoro se očekuje i dobivanje utrke za muškarce." (...)

bile otišle. Nazvala sam _____ muža da dođe po _____. Rekao je da će _____, naravno, ali tek _____ završi sastanak.

Riječ je o zadacima koji se često upotrebljavaju u *cloze-testovima* iz engleskog jezika kod kojih se sustavno ispušta svaka 5. ili 7. riječ. Za engleski jezik takav tip zadatka za dopunjavanje očito djeluje, međutim u slavenskim je jezicima zbog većinom bogate fleksijske strukture problematičan. Iako je zadatak u (10) razmjerno lako pravilno rješiv, ima neke nedostatke. Nije naime jasno što je s nelogičnim ili nepravilno zapisanim odgovorima (uputa ne zahtijeva logičnost ili pravilnost), koliko riječi treba upisati na svaku crtu itd. Kod tog se zadatka dakle otvara pitanje njegove valjanosti. Kako bi npr. bio ocijenjen ovakav inovativan odgovor: "Nazvala sam *njezinog* muža da dođe *poslijepodne*"? Budući da vjerojatno nije bio predviđen, za njega nećemo dobiti nijedan bod, iako je takav odgovor i te kako moguć. Autori jezičnih testova u takvim su slučajevima u donekle paradoksalnoj situaciji: budući da jedini poznaju pravilni odgovor, samo o njihovoj toleranciji ovisi hoće li se prihvatiti i drukčiji — vjerojatni i prihvatljivi odgovori. Da bi se izbjegle nedoumice kod ocjenjivanja, predlažemo da se u takvom zadatku ispuste samo određene riječi koje imaju neki zajednički nazivnik: možda veznici, prijedlozi ili zamjenice, posvojne zamjenice ili slično, kao u primjeru u (11).

- (11) *Popunite tekst prijedlozima kao što je pokazano u primjeru. (1 bod)*

Ne odugovlačite (0) s odlukama, kućnim poslovima ili pripremama (1) _____ izlazak. Neočekivani poziv brzo će vas razbuditi, a radi se (2) _____ nekome tko vam se jako sviđa. Ukućani će tražiti (3) _____ vas da ispunite svoje obaveze.

Zadatak možemo ispitanicima donekle olakšati ako im na izbor ponudimo niz riječi ili sintagmi kojima moraju smisleno i pravilno dopuniti tekst. I u tom primjeru treba ponuditi više riječi nego što ih nedostaje, kao u (12).

- (12) *Upišite u tekst riječi koje nedostaju, a ponuđene su na kraju teksta. Pazite, treba ih staviti u odgovarajući oblik. U popisu su dvije riječi suvišne i ne pripadaju ovome tekstu.*

Dionice Croatia osiguranja od ponedjeljka u prvoj (0) *kotaciji* Zagrebačke burze

ZAGREB, 8. (1) _____ — Predsjednik Uprave Croatia osiguranja Marijan Ćurković i direktor Zagrebačke burze Marinko Papuga u ponedjeljak će potpisati ugovor o uvrštenju (2) _____ Croatia osiguranja

u prvu kotaciju Zagrebačke burze. Trgovanje redovnim i (3) _____ dionicama Croatije dosad se trgovalo u JDD kotaciji, a u utorak, 13. siječnja, počinje trgovanje u prvoj kotaciji. (...)

redovan fond povlašten siječanj dionica

Taj tip zadatka spada u zadatke dopunjavanja, pri čemu se kandidatima nudi popis riječi ili fraza koje može izabrati. Kod njega ispitanici moraju “razbijeni” tekst sastaviti u koherentnu cjelinu. Budući da je i tu riječ o izboru, treba paziti na ponuđene mogućnosti jer smisljena i pravilna može uvijek biti samo jedna jedina mogućnost. I tu obično dodajemo distraktor ili dva da se zadatak oteža. Ali pazite: distraktori ni u kom slučaju ne smiju odgovarati, ali moraju biti vjerojatni jer će ih samo tako netko izabrati.

3.5. Zadaci poluotvorenoga tipa

(13) *Pročitajte tekst na sljedećoj stranici i odgovorite na ova pitanja.*

Primjer: Koja mogućnost rehabilitacije odnedavno postoji za osobe s invalidnošću?

Veslanje je nova mogućnost rehabilitacije.

Je li veslanje za osobe s invalidnošću u Hrvatskoj razvijeno?

Premda je Hrvatska bila prisutna na posljednjem Svjetskom veslačkom prvenstvu u Etonu 2006. s dvije posade veslača s invalidnošću veslanje za osobe s invalidnošću u Hrvatskoj za sada je dobro razvijeno samo u Zagrebu, i to najviše zahvaljujući entuzijizmu stručnog vodstva veslačkog kluba “Sava”. U drugim dijelovima Hrvatske tek je u začetku...

Zadaci s pitanjima koji zahtijevaju kratke ili dulje odgovore spadaju u najautentičnije. Ispitanici ih većinom zato vrlo dobro rješavaju; tako naime pokazuju statističke analize slovenskih testova. Pa ipak odlomak u (13) pokazuje da zadatak nije pravilno sastavljen. Najprije je tu primjer koji zbunjuje (čak je nepravilan). Odgovor *Veslanje* bio bi posve dovoljan i to bi za ispitanike bio znak da napišu ono bitno. Odlučujuća pitanja ne pripadaju u takav tip zadataka. Za odgovore *da* i *ne* mnogo je prikladniji prvi tip, kao u (5). A što biste učinili kad bi ispitanik odgovorio znakom +? Zato najprije moramo dati uputu u kojoj će biti objašnjeno koliko približno riječi treba sadržavati odgovor. Zatim predložimo oblikovanje dopunskih pitanja: tko, što, kada, gdje, zašto. Odgovorima na njih ispitanici će dokazati jesu li

tekst shvatili ili ne, odnosno koliko su u njemu razumjeli. Pitanja također moraju biti napisana ispred teksta (odnosno pročitana unaprijed ako ispitanici slušaju tekst). Ako znamo koju informaciju tražimo, lakše ćemo je naći. U tom zadatku treba upozoriti na još nešto. Ako njime mjerimo razumijevanje, moramo u odgovorima dopustiti gramatičke i pravopisne pogreške (naravno, ako one ne utječu na razumijevanje). Ako to ne učinimo te odbijemo bodove i za gramatičke pogreške, prekršit ćemo jedan od najvažnijih zahtjeva — zahtjev za valjanošću.

3.6. Zadaci otvorenoga tipa

Otvoreni tipovi zadataka nekako su rezervirani za govornu i pisanu produkciju. Tu međutim nisu problematični zadaci, nego ocjenjivanje. O ocjenjivanju nekoliko kratkih riječi na kraju ovoga rada, a ovdje još dva primjera:

- (14) *Izaberite jednu temu i napišite tekst.*
 Pojedinaac u suvremenom društvu
 Najdraži film
 Da sam političar

Dobro je da ispitanici imaju mogućnost izbora. Međutim, problem je kod zadatka u (14) upravo u tome: ponuđene mogućnosti međusobno nisu usporedive. Ako se nude naslovi koji su možda primjereni za slobodne sastave u školi, odrasli, koji uglavnom pišu praktične tekstove, teško će se identificirati s njima. Zato odraslima radije ponudimo primjerenije vrste tekstova (koji, s druge strane, nisu primjereni za djecu): zamolbu, žalbu, životopis, ispriku itd. Zadatak u (14) nedostatan je u još nečemu: nije navedeno koliko dug mora biti tekst, koje elemente treba sadržavati i što će se ocjenjivati. I zato će biti teško uspoređivati različita rješenja. Kao primjer dobre prakse predložimo zadatak u (15).

- (15) *Svom susjedu napišite razglednicu s putovanja. Neka tekst sadržava oko 30 riječi. Napišite gdje ste, što ste već vidjeli, kakve još planove imate, što vam se sviđa ili ne sviđa... Sve podatke možete izmisliti. Nemojte zaboraviti uputiti srdačan pozdrav. Pazite na sadržaj i jezičnu pravilnost.*

Kao što je već rečeno: problem produktivnih zadataka nije toliko pitanje samih zadataka. Veće poteškoće pritom uzrokuje ocjenjivanje rješenja koja nastaju na temelju njih. Ako je testove objektivnoga tipa razmjerno lako bodovati (s obzirom na bodovnu šablonu koju treba napraviti za svaki test, po jedan bod obično dodjeljujemo za svaki pravilni odgovor), kod tzv. subjek-

tivnog ocjenjivanja sve se komplicira. Izrada preciznih i jasnih ocjenjivačkih ljestvica spada u fazu planiranja testa — u specifikaciju testa.

Ovaj kratki zapis ne dopušta temeljitije razmatranje načina bodovanja i ocjenjivanja. Problematika ocjenjivanja produktivnih dijelova testova vrlo je složena i zahtijevala bi samostalnu obradu; konačno, o njoj su napisane stotine knjiga. Zato želimo zaključiti mišlju o odgovornosti koju za svoj rad moraju preuzeti ispitivači.

4. Zaključak

Iz svega rečenoga možemo zaključiti da izrada testova mora biti dobro promišljena jer testovi uvijek odlučuju o pojedincima — više-manje sudbonosno. Upravo o rezultatima na testu, točnije o tome kako se interpretiraju, često ovisi hoće li neka osoba napredovati ili ne, hoće li biti “prihvaćena” ili “odbijena”. Testiranje je zato u svim svojim fazama — od planiranja, izrade, izvedbe testiranja, ocjenjivanja do interpretacije rezultata testova — velika odgovornost. Pri planiranju bilo kojega jezičnog testa zato treba razmišljati i o posljedicama koje test može imati, pogotovo kod loših rezultata. Za loš rezultat može biti puno razloga, ne samo individualnih. Pokazali smo da na rezultate u testu može utjecati sam test: nepravilno izabrani i pripremljeni zadaci, (nepravilne) upute uz zadatke i konačno tekstovi na kojima se zadaci zasnivaju. Zato rezultati testova uvijek trebaju biti povratna informacija i za autore testa — informacija o testu, a ako je on dio nastavnoga procesa, i o nastavnom procesu. Osim toga polazište svakoga testa uvijek mora biti njegov neposredni korisnik — ispitanik, u našem slučaju to je korisnik jezika (prvoga ili neprvoga). Postoje i drugi korisnici, kao što su škole koje provode jezične testove i njihovi učitelji koji na ispite pripremaju, poslodavci i zakonodavci koji oblikuju zahtjeve za jezično znanje, obitelji ispitanika itd.

Jezično testiranje složena je tematika koja se ne može iscrpiti jednim jedinim člankom. Zato je na ovome mjestu bilo moguće istaknuti samo neka pitanja, sigurno važna, ali još uvijek ne iscrpna. Tako bi npr. samostalnu obradu zahtijevalo ocjenjivanje, posebna bi se pozornost morala posvetiti analizi i interpretaciji rezultata itd.; testiranje govorne produkcije na ovome mjestu nismo ni spomenuli. A svrha je ovoga rada bila u prvome redu potaknuti sve koji se bave jezičnim testovima ili se njima tek namjeravaju baviti na stručno i etično promišljanje.

5. Literatura

- Bachman, L. F. (1990) *Fundamental Considerations in Language Testing*, Oxford: Oxford University Press.
- Bachman, L. F. i Palmer, S. (1996) *Language Testing in practice*, Oxford: Oxford university Press.
- Ferbežar, I. (1999) Merjenje in merljivost v jeziku. Na stičišču jezikoslovja in psihologije: nekaj razmislekov, *Slavistična revija* 47/4, 418–436.
- Ferbežar, I. i Pirih-Svetina, N. (2002) *Izobraževalni programi, Izobraževanje odraslih, Slovenščina za tujce*, Ljubljana: Ministrstvo za šolstvo, znanost in šport: Zavod RS za šolstvo, 44 str.
www.acs.si/programoteka
- Ferbežar, I. i sur. (2004) Slovenski pojmovnik s področja jezikovnega testiranja, u Martyniuk, W. i Ferbežar, I. (ur.) *English-Polish-Slovene glossary of language testing terms*, (Opisywanie i testowanie biegłości językowej, 3), Kraków TAIWPN Universitas, 115–167.
- Ferbežar, I. i Bajec, B. (2004) *Priročnik za avtorje testnih gradiv* (Opisywanie i testowanie biegłości językowej, 10), Ljubljana: Center za slovenščino kot drugi/tuji jezik — Kraków: TAIWPN Universitas.
- Požgaj-Hadžić, V. i Balažić-Bulc, T. (2004) *Hrvatski jezik, Srpski jezik, Vodič za polaganje ispita na visokoj razini*, Ljubljana: Oddelek za slavistiko Filozofske fakultete u Ljubljani.
- Spillner, B. (1995) Verstehen und Verständlichkeit: die Rezeptionsseite sprachlicher Kommunikation, u Spillner, B. (ur.) *Sprache: Verstehen und Verständlichkeit. Kongressbeiträge zur 25. Jahrestagung der Gesellschaft für Angewandte Linguistik GAL e. V.* 13–14, Frankfurt am Main — Berlin — Bern — New York — Paris — Wien: Peter Lang.

Designing language tests

Measuring language proficiency in learning/teaching a foreign/second language is a topical issue of European language policies. The most reliable (formal) measuring instruments are language tests, which must satisfy the defined test-design criteria. In the case of the communicative method of teaching and measuring knowledge, the principal concept of test design and delivery is authenticity — not only text authenticity but also authenticity of test situations. Different characteristics of tests are described in the paper (from validity as most important to objectivity, sensitivity, etc.) as well as different purposes of testing (from informal measuring of teaching efficiency to the formal public system of language proficiency certification). Special attention is paid to the many open problems of test design, such as selection of texts, different types of test items (and whether they should explicitly test grammar), grading, that is, grading and evaluation criteria etc. On examples of Croatian language tests, numerous “pitfalls” in test design

are highlighted, which test designers should keep in mind. It is concluded that test design should be well thought-out, for test results may sometimes have fateful consequences for an individual. Therefore, test designers and deliverers must be aware of their responsibility.

Key words: Croatian as a foreign language, language testing, characteristics of tests, different types of test items

Ključne riječi: hrvatski jezik kao strani, jezično testiranje, izrada i karakteristike testova, tipovi zadataka